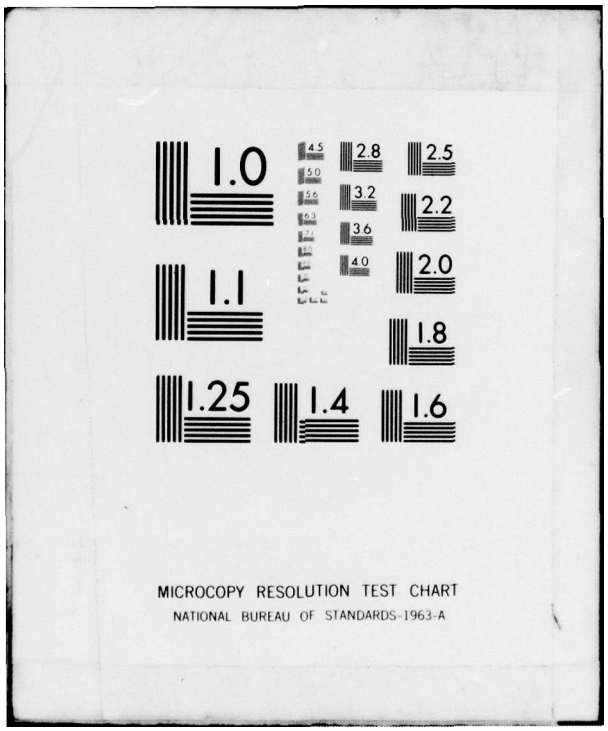
AD-A077 097 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
QUASI-NEWTON METHODS FOR GENERALIZED EQUATIONS,(U)
JUN 79 N H JOSEPHY DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1966 NL

| OF |
AD
A077097

END
DATE
FILMED
12-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
8.0
9.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(10)

AD A077092

MRC Technical Summary Report #1966

QUASI-NEWTON METHODS FOR GENERALIZED EQUATIONS

Norman H. Josephy

LEVEL II

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

DDC
RECEIVED
NOV 26 1979
E

DDC FILE COPY

June 1979

Received May 25, 1979

See 1473 in back

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

QUASI-NEWTON METHODS FOR GENERALIZED EQUATIONS

Norman H. Josephy

Technical Summary Report #1966
June 1979

ABSTRACT

Newton's method is a well known and often applied technique for computing a zero of a nonlinear function. Situations arise in which it is undesirable to evaluate, at each iteration, the derivative appearing in the Newton iteration formula. In these cases, a technique of much modern interest is the quasi-Newton method, in which an approximation to the derivative is used in place of the derivative. By using the theory of generalized equations, quasi-Newton methods are developed to solve problems arising in both mathematical programming and mathematical economics.

We prove two results concerning the convergence and convergence rate of quasi-Newton methods for generalized equations.

We present computational results of quasi-Newton methods applied to a nonlinear complementarity problem of Kojima [11].

AMS(MOS) Subject Classification: 49D99, 90C99, 90A15

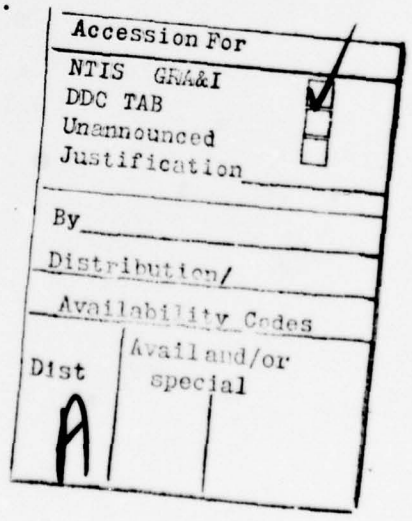
Key Words: Variational Inequalities, Complementarity, Ouasi-Newton Methods.

Work Unit Number 5 - Mathematical Programming and Operations Research

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024. This material is based upon work supported by the National Science Foundation under Grant No. DCR 74-20584 and Grant No. MCS 74-20584 A02 and the Graduate School of the University of Wisconsin - Madison.

Significance and Explanation

Many practical problems in operations research and mathematical economics can be formulated as a system of equations and inequalities. A previous paper (Josephy [10]) developed an iterative procedure, Newton's method, for computing a solution to such a system. However, situations arise in which it is computationally unreasonable to evaluate the derivative appearing in the Newton iteration formula. In such cases, an approximation to the derivative which is easier to compute is used in place of the derivative. The iterative procedure with an approximation to the derivative is called a quasi-Newton method. This paper extends the results known about the convergence and rate of convergence of quasi-Newton methods for equations to the case of equations and inequalities which arise in the nonlinear programming problem and the economic equilibrium problem. The method is illustrated by solving a small practical problem involving equations and inequalities.



The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

QUASI-NEWTON METHODS FOR GENERALIZED EQUATIONS

Norman H. Josephy

1. Introduction.

We recall the definition of a generalized equation. Further elaboration can be found in Robinson [13-17] and Josephy [10].

Let C be a non-empty, convex, closed subset of R^n. The normal cone to C at $x \in C$ is the set of outward pointing normals to C at x. We have the following:

Definition 1. Let C be a non-empty, closed, convex subset of R^n.

The normal cone to C at x is given by

$$N_C(x) := \begin{cases} \{z \mid \langle z, k-x \rangle \leq 0 \ \forall \ k \in C\} & \text{if } x \in C, \\ \phi & \text{if } x \notin C \end{cases}$$

We can now define a generalized equation.

Definition 2. Let $f: D \subset R^n \to R^n$.

Let C be a non-empty, closed, convex subset of R^n.

A generalized equation is a set relation

$$0 \in f(x) + N_C(x) . \tag{1}$$

Thus, x^* satisfies the generalized equation $0 \in f(x) + N_C(x)$ if and only if x^* satisfies the relations

$$x^* \in C$$

and

$$\langle f(x^*), k - x^* \rangle \geq 0 \quad \text{for all } k \in C .$$

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024. This material is based upon work supported by the National Science Foundation under Grant No. DCR 74-20584 and Grant No. MCS 74-20584 A02 and the Graduate School of the University of Wisconsin - Madison.

Newton's method for solving a generalized equation $0 \in F(x) + N_C(x)$ involves the evaluation of F', the derivative of F, at each iterate x_n. (See Josephy [10]). In situations where this evaluation imposes an excessive computational burden, we consider replacing $F'(x_n)$ with an approximation. A quasi-Newton method for solving $0 \in F(x) + N_C(x)$ replaces F with an affine map $F(x_n) + B(\cdot - x_n)$ whose derivative B is both an approximation to $F'(x_n)$ in some sense and is computationally easier to evaluate. One class of approximations, the secant approximations, has been the subject of great interest to those solving systems of equations and related problems, such as non-linear least squares estimation, unconstrained and equality-constrained optimization, and, most recently, inequality-constrained optimization. A sample of the recent work on quasi-Newton methods can be found in Brodlie [2], Dennis and Moré [5], Gill and Murray [8], Tapia [18], and Han [9].

In this paper, we will extend to the generalized equation case two fundamental results on the convergence and rate of convergence of quasi-Newton methods with secant approximations. Section 2 contains introductory material on quasi-Newton methods. We prove local convergence of quasi-Newton iterates to a solution of a generalized equation in Section 3, and prove in Section 4 Q-superlinear rate of convergence for the quasi-Newton iterates. We conclude in Section 5 with the computational results of applying a quasi-Newton method to the generalized equation representing a nonlinear complementarity problem of Kojima [11].

2. Quasi-Newton Methods.

Newton's method for finding a zero of F proceeds by iteratively solving $0 = F(x_n) + F'(x_n)(x-x_n)$ for a solution x_{n+1}. A quasi-Newton method iteratively solves $0 = F(x_n) + B_n(x-x_n)$, where B_n approximates $F'(x_n)$ in some sense. An approximation of much current interest is a generalization to maps on R^n of the secant approximation of the derivative of a real-valued map on R. Let $f:R \to R$ have derivative f'. Given two points in R, say x_n and x_{n+1}, the secant approximation to $f'(x_{n+1})$ is

$$b := (f(x_{n+1}) - f(x_n))/(x_{n+1} - x_n) .$$

Alternatively, the secant approximation b is the unique solution of

$$f(x_{n+1}) - f(x_n) = b(x_{n+1} - x_n) . \tag{1}$$

For $F:R^n \to R^n$, a secant approximation to $F'(x_{n+1})$ is any matrix B satisfying

$$F(x_{n+1}) - F(x_n) = B(x_{n+1} - x_n) . \tag{2}$$

Since equation (2) does not uniquely specify B, additional conditions can be imposed to guarantee desired properties of B, such as symmetry and positive definiteness. This approach to secant approximations is discussed in more detail in Dennis and Schnabel [6] . Stable numerical techniques for efficiently implementing these secant approximations are described in Gill, Golub, Murray and Saunders [7]. The price one pays for using an approximation for $F'(x_n)$ is the

loss of the quadratic rate of convergence to a local solution. However, iterates generated by secant approximations typically exhibit Q-superlinear convergence, provided appropriate conditions are satisfied. We will need the concepts of linear and Q-superlinear convergence, which we now define.

Definition 1. A sequence $\{x_n\}$ converges to x^* linearly (Q-superlinearly) if and only if for some $a \in (0,1)$ (for some sequence $\{a_n\}$ converging to zero),

$$\| x_{n+1} - x^* \| \leq a \| x_n - x^* \|$$

$$(\| x_{n+1} - x^* \| \leq a_n \| x_n - x^* \|) .$$

The Q-superlinear convergence of quasi-Newton iterates can be proven in the following fashion. In Theorem 1, we establish linear convergence for iterates determined by any quasi-Newton method whose approximations $\{B_n\}$ satisfy a certain norm inequality. The proof for the generalized equation case is modeled on the single-valued case given in Broyden, Dennis and Moré [1, Theorem 3.2]. We can then invoke the results appearing in Broyden, Dennis and Moré [1] which show that the traditional update formulas, such as the Broyden rank one, Powell symmetric Broyden rank two, and the DFP update of $\{B_n\}$ (see Dennis and Moré [5] for further details) satisfy the required norm inequality. This establishes a linear convergence rate for the generalized equation case. Dennis and Moré [4, Theorem 3.4 and Section 4] show that

$$\ell := \lim_{n \to \infty} \| (B_n - F'(x^*))(x_{n+1} - x_n) \| / \| x_{n+1} - x_n \| = 0 ,$$

whenever the sequence $\{B_n\}$ is computed by one of the above mentioned update formulas and the sequence $\{x_n\}$ converges linearly to x^*. They note, in a remark immediately following their proof of Theorem 3.4, that the limit $\ell=0$ is established independent of how the sequence $\{x_n\}$ is generated. In particular, their result remains valid when $\{x_n\}$ is the linearly convergent sequence of solutions to the quasi-Newton method applied to generalized equations rather than to single-valued equations. Thus, for the update formulas mentioned above, the limit $\ell=0$ is valid for the generalized equation case, since Theorem 1 of this paper establishes the linear convergence of $\{x_n\}$. Theorem 2 shows that $\ell=0$ suffices to establish Q-superlinear convergence of iterates $\{x_n\}$ generated by quasi-Newton methods applied to generalized equations. We note that proofs of the two theorems appearing in this paper, when restricted to the single-valued equation case, have appeared in the literature (Dennis and Moré [5]). However, those proofs depend critically upon the fact that single-valued equations are being solved. Hence, proofs are given in this paper which are valid for the generalized equation case.

We conclude this section with a corollary to Theorem 2.4 of Robinson [16], and a lemma from Ortega and Rheinboldt [12].

Corollary 1. Let C be a closed, convex, nonempty subset of R^n, and let D be an open, convex, nonempty subset of R^n. Let $f:D \to R^n$ have Fréchet derivative f'. Suppose the generalized equation $0 \in f(x) + N_C(x)$ has a strongly regular solution at $x^* \in D$, with associated Lipschitz constant d.

Then, for some positive constants r, R, b and e, the following hold.

> Let A be an $n \times n$ matrix and let $\bar{x} \in R^n$. Then
> $\bar{B}(x^*,r) \cap (f(\bar{x}) + A\cdot(\cdot-\bar{x}) + N_C)^{-1}$ restricted to $\bar{B}(0,R)$
> is single-valued and Lipschitz continuous with modulus
> $d(1-d\,\|A-f'(x^*)\|)^{-1}$, whenever $\|\bar{x}-x^*\| < e$ and
> $\|A-f'(x^*)\| < b$.

Lemma 1. Let D be an open, convex subset of R^n. Let $f:D \to R^n$ be continuously differentiable. Suppose that for some $k > 0$, $\|f'(u) - f'(v)\| \leq k\|u - v\|$ whenever $u,v \in D$. Then $\|f(u) - Lf_v(u)\| \leq \frac{1}{2} k\|u - v\|^2$ whenever $u,v \in D$, where $Lf_v(u) := f(v) + f'(v)(u-v)$.

The definition and properties of strong regularity of a solution to a generalized equation can be found in Robinson [16]. The use of strong regularity in establishing convergence properties of Newton's method for generalized equations can be found in Josephy [10].

3. Local Convergence.

We begin with a definition.

Definition 2. (Dennis and Moré [5]).

Let D be an open, nonempty convex subset of R^n, and let P denote a class of $n \times n$ matrices. Then an update function U is a map from $D \times P$ to subsets of P.

An update function can be used to define a set of approximations to the derivative of F at x_{n+1}. In the typical situation, the present iterate x_n and the present approximation B_n will determine the new iterate x_{n+1}. Both iterates x_n and x_{n+1} are used to determine the next approximation B_{n+1}. This procedure, by which B_{n+1} is determined from x_n and B_n, can be represented by $B_{n+1} \in U(x_n, B_n)$, where U is the appropriately defined update function.

We now state and prove the local convergence theorem.

Theorem 1. Let D be an open, nonempty convex subset of R^n, and let C be a closed, nonempty convex subset of R^n. Let $F: D \to R^n$ have a Lipschitz continuous derivative F' with Lipschitz constant K. Suppose $x^* \in D$ is a strongly regular solution of the generalized equation $0 \in F(x) + N_C(x)$ with associate Lipschitz constant d. Let W be an open neighborhood of $F'(x^*)$ in the space of linear maps from

R^n to R^n. Let $\|\cdot\|_M$ denote a matrix norm and let $a > 0$ be such that $\|\cdot\| \leq a\|\cdot\|_M$, where $\|\cdot\|$ is a matrix norm subordinate to the given vector norm on R^n. Suppose that two positive constants, a_1 and a_2, exist such that the inequality

$$\|\bar{B} - F'(x^*)\|_M \leq (1+a_1 \cdot \max\{\|\bar{x}-x^*\|, \|x-x^*\|\})\|B-F'(x^*)\|_M + a_2 \cdot \max\{\|\bar{x}-x^*\|, \|x-x^*\|\}$$

holds whenever the following conditions are valid:

$(x,B) \in D \times W$, $\bar{B} \in U(x,B)$, where U is an update function, and $\bar{x}$, the vector closest to x in the set $(F(x) + B(\cdot - x) + N_C)^{-1}(0)$, exists.

Let b, e, r and R be the positive constants associated with the strongly regular solution x^*, as given in Corollary 1. Fix p positive and less than 1. Suppose the following relations hold, with e and b reduced, if necessary, from those guaranteed by Corollary 1.

$$\|B_0 - F'(x^*)\|_M < \tfrac{1}{2}ba^{-1} \quad \text{and} \quad \|x_0 - x^*\| < e.$$

$$(\tfrac{3}{2}Ke+b)e < R \quad \text{and} \quad d(1-db)^{-1}(\tfrac{3}{2}Ke+b) < p.$$

$$2a(a^{-1}ba_1+a_2)e\,(1-p)^{-1} < b \quad \text{and} \quad \bar{B}(x^*,e) \subset D.$$

Then a sequence of iterates $\{x_n\}$ and a sequence of matrices $\{B_n\}$ exist and satisfy the following relations, for all $n \geq 0$.

$$0 \in F(x_n) + B_n(x_{n+1}-x_n) + N_C(x_n) \quad .$$

$$B_{n+1} \in U(x_n, B_n) \quad .$$

$$\| B_n - F'(x^*) \|_M < ba^{-1} \quad .$$

$$\| x_{n+1}-x^* \| \leq p \| x_n - x^* \| \quad .$$

Proof. We will use the following notation.

$$S_n := \bar{B}(x^*,r) \cap (F(x^*) + B_n(\cdot - x^*) + N_C)^{-1} \quad .$$

$$T_n := \bar{B}(x^*,r) \cap (F(x_n) + B_n(\cdot - x_n) + N_C)^{-1} \quad .$$

$$J_n := F(x^*) - F(x_n) - B_n(x^* - x_n) \quad .$$

$$\bar{d} := d(1-db)^{-1} \quad .$$

By hypothesis, $\| B_0 - F'(x^*) \|_M < \frac{1}{2}ba^{-1}$ and $\| x_0 - x^* \| < e$. Thus, $\| B_0 - F'(x^*) \| < b$. By Corollary 1, T_0 restricted to $\bar{B}(0,R)$ is single-valued and Lipschitz continuous with Lipschitz constant $d(1-d\| B_0 - F'(x^*) \|)^{-1}$, which is less than $\bar{d}$. Similarly, S_0 restricted to $\bar{B}(0,R)$ is single-valued and Lipschitz continuous with Lipschitz constant $\bar{d}$, and $S_0(0) = x^*$. Define $x_1 := T_0(0)$. In order to estimate the distance from x_1 to x^*, we need to estimate the norm of J_0. It follows from

$$\begin{aligned} J_0 &:= F(x^*) - F(x_0) - B_0(x^* - x_0) \\ &= F(x^*) - LF_{x_0}(x^*) + (F'(x^*) - B_0)(x^* - x_0) + \\ &\quad + F'(x_0) - F'(x^*)(x^* - x_0) \end{aligned}$$

that the norm of J_0 is bounded by

$$\| J_0 \| \leq \| F(x^*) - LF_{x_0}(x^*) \| + \| F'(x^*) - B_0 \| \cdot \| x^* - x_0 \| +$$

$$+ \| F'(x^*) - F'(x_0) \| \cdot \| x^* - x_0 \|$$

$$\leq \tfrac{1}{2} K \| x^* - x_0 \|^2 + b \| x^* - x_0 \| + K \| x^* - x_0 \|^2$$

$$\leq (\tfrac{3}{2} Ke + b) \| x^* - x_0 \| < R .$$

By definition of T_0, $x_1 \in \bar{B}(x^*; r)$ and

$$0 \in F(x_0) + B_0(x_1 - x_0) + N_C(x_1)$$

$$= F(x^*) + B_0(x_1 - x^*) - J_0 + N_C(x_1)$$

Thus, $J_0 \in F(x^*) + B_0(x_1 - x^*) + N_C(x_1)$, so that $x_1 \in S_0(J_0)$. But $\| J_0 \| < R$, implying $S_0(J_0)$ is a singleton and $x_1 = S_0(J_0)$. We can now estimate the distance between x_1 and x^* as

$$\| x_1 - x^* \| = \| S_0(J_0) - S_0(0) \| \leq \bar{d} \| J_0 \|$$

$$\leq \bar{d} (\tfrac{3}{2} Ke + b) \| x^* - x_0 \| < p \| x_0 - x^* \| .$$

We now proceed by induction. We will show that for all $k \geq 0$,

$$\| B_k - F'(x^*) \|_M < ba^{-1} , \tag{3}$$

$$\| x_{k+1} - x^* \| \leq p \| x_k - x^* \| \leq p^k \| x_0 - x^* \| . \tag{4}$$

We have already established (3) and (4) when k=0. We thus assume that (3) and (4) hold for all $k \leq m-1$, where $m \geq 1$, and will establish

(3) and (4) for k=m. We first note that $p < 1$ and (4) imply

$$m_k := \max\{\| x_{k+1}-x^* \| , \| x_k-x^* \|\} = \|x_k-x^*\|$$

for all $k \leq m-1$, and

$$\| B_{k+1}-F'(x^*) \|_M - \| B_k-F'(x^*) \|_M \leq a_1 m_k \| B_k-F'(x^*) \|_M + a_2 m_k$$

$$\leq a_1 b a^{-1} p^k \| x_0-x^* \| + a_2 p^k \| x_0-x^* \|$$

$$\leq (a_1 b a^{-1} + a_2) p^k e \quad .$$

Summing over k between 1 and m-1 yields

$$\| B_m-F'(x^*) \|_M \leq \|B_0-F'(x^*) \|_M + (a_1 b a^{-1} + a_2)(1-p)^{-1} e \quad .$$

By hypothesis, $\|B_0-F'(x^*) \|_M < \frac{1}{2} b a^{-1}$ and $(a_1 b a^{-1}+a_2)(1-p)^{-1} e < \frac{1}{2} b a^{-1}$, hence $\| B_m-F'(x^*) \|_M < b a^{-1}$. Thus (3) is established. To prove (4), we note that $\| x_m-x^* \| \leq p^m \| x_0-x^* \| \leq p^m e < e$. We can now apply Corollary 1 to T_m and S_m and conclude that, when restricted to $\bar{B}(0,R)$, both are single-valued and Lipschitz continuous with Lipschitz constant $d(1-d\| B_m-F'(x^*) \|)^{-1}$. Let $x_{m+1} := T_m(0)$ and note that $x^* = S_m(0)$. It remains to bound the distance between x_{m+1} and x^*. We will first obtain a bound on the norm of J_m, and then use that bound to estimate $\|x_{m+1}-x^*\|$. We have that

$$J_m := F(x^*) - F(x_m) - B_m(x^*-x_m)$$

$$= (F(x^*) - LF_{x_m}(x^*)) + (F'(x^*) - B_m)(x^*-x_m) +$$

$$+ (F'(x_m) - F'(x^*)(x^*-x_m) .$$

Using Lemma 1 to bound the first term, (3) to bound the second term, and Lipschitz continuity of F' to bound the third term, we have

$$\| J_m \| \leq \tfrac{1}{2}K \| x_m-x^* \|^2 + b \| x_m-x^* \| + K \| x^*-x_m \|^2$$

$$\leq (\tfrac{3}{2}Ke+b) \| x_m-x^* \| < R .$$

By definition of T_m, $x_{m+1} \in \bar{B}(x^*,r)$ and

$$0 \in F(x_m) + B_m(x_{m+1}-x_m) + N_C(x_{m+1})$$

$$= F(x^*) + B_m(x_{m+1}-x^*) - J_m + N_C(x_{m+1}) .$$

Hence, $x_{m+1} \in S_m(J_m)$ and $\|J_m\| < R$, from which we can conclude that $x_{m+1} = S_m(J_m)$. We finish the induction proof by noting that

$$\| x_{m+1}-x^* \| = \| S_m(J_m) - S_m(0) \|$$

$$\leq d(1-d \| B_m - F'(x^*) \|)^{-1} \| J_m \|$$

$$\leq \bar{d}(\tfrac{3}{2}Ke+b) \| x_m-x^* \|$$

$$< p \| x_m-x^* \| .$$

This completes the induction and the proof of the theorem.

4. Q-Superlinear Convergence

The preceding theorem establishes local linear convergence of the quasi-Newton iterates $\{x_n\}$ to a strongly regular solution x^* of the generalized equation, for any update function which satisfies the given norm inequality. The results of Broyden, Dennis and Moré [1] show that the traditional update functions satisfy this inequality. Dennis and More [4] show that for the traditional update functions, the limit appearing in the next theorem is zero. Theorem 2 proves that this suffices for the quasi-Newton iterates to converge Q-superlinearly.

Theorem 2. Let D be an open, nonempty convex subset of R^n, and let C be a closed, nonempty convex subset of R^n. Let F have a Lipschitz continuous derivative F' with Lipschitz constant K. Suppose that the generalized equation $0 \in F(x) + N_C(x)$ has a strongly regular solution x^* with associated Lipschitz constant d. Let $\{B_k\}$ be a sequence of $n \times n$ matrices. Assume that the set of norms $\{\|B_k\|\}$ is bounded. Let $x_0 \in D$ and suppose that the sequence $\{x_k\}$, $k \geq 0$ satisfies the relation $0 \in F(x_k) + B_k(x_{k+1}-x_k) + N_C(x_{k+1})$. Also assume that $\{x_k\}$ converges to x^*. Define

$$E_k := B_k - F'(x^*), \quad s_k := x_{k+1}-x_k, \text{ and } v_{k+1} := -F(x_k) - B_k s_k .$$

Then $\lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0$ implies that the sequence $\{x_k\}$ converges

Q-superlinearly to x^*.

Proof. By definition and some algebra, we have

$$E_k s_k = F(x_{k+1}) - F(x_k) - F'(x^*)s_k - (v_{k+1} + F(x_{k+1})) .$$

Letting $p_{k+1} := F(x_{k+1}) + v_{k+1}$, we can solve for p_{k+1} and take norms to obtain

$$\frac{\| p_{k+1} \|}{\| s_k \|} \leq \frac{\| \Delta F \|}{\| s_k \|} + \frac{\| E_k s_k \|}{\| s_k \|} ,$$

where $\Delta F := F(x_{k+1}) - F(x_k) - F'(x^*)(x_{k+1}-x_k)$. But the Lipschitz continuity of F' implies (see Ortega and Rheinboldt [12]) $\| \Delta F \| \leq K \cdot \max\{ \| x_{k+1}-x^* \| , \| x_k-x^* \| \} \cdot \| x_{k+1}-x_k \|$. Hence,

$$\lim_{k\to\infty} \frac{\| p_{k+1} \|}{\| s_k \|} \leq \lim_{k\to\infty} K \cdot \max \{ \| x_{k+1}-x^* \| , \| x_k-x^* \| \} +$$

$$+ \lim_{k\to\infty} \frac{\| E_k s_k \|}{\| s_k \|}$$

$$= 0 .$$

To obtain the next result, we use a special case of the Implicit Function Theorem of Robinson (16 , Theorem 2.1). Specifically, we take as the function f(p,x) in that theorem the function -p +F(x) . The conclusions of that theorem give us the following results. Fixing $e > 0$, there exist neighborhoods U_e of 0 and W_e of x^* , and a single-valued map $\bar{x}: U_e \to W_e$ such that $\bar{x}(p)$ is the unique solution in W_e

of the generalized equation $0 \in -p + F(x) + N_C(x)$. Also, for any $p, q \in U_e$, $\| \bar{x}(p) - \bar{x}(q) \| \leq (d+e) \| p+q \|$. We now show that this result can be applied to the generalized equation $0 \in F(x_k)+B_k s_k + N_C(x_{k+1})$, where x_{k+1} is the unknown variable, to obtain a bound on $\| x_{k+1} - x^* \|$. Note that the assumptions that $\{\| B_k \|\}$ is bounded and $\{x_k\}$ converges to x^* imply the convergence of $\{v_{k+1}\}$ to $-F(x^*)$. Thus, $\{p_{k+1}\}$ converges to zero, and will be in U_e for all sufficiently large k. Also, $\bar{x}(0) = x^*$. Thus, $x_{k+1} \in W_e$,

$0 \in F(x_k) + B_k s_k + N_C(x_{k+1}) = F(x_{k+1}) - p_{k+1} + N_C(x_{k+1})$ and

$p_{k+1} \in U_e$ for all sufficiently large k, which implies $\bar{x}(p_{k+1}) = x_{k+1}$ for all sufficiently large k. Hence,

$$\| x^*-x_{k+1} \| = \| \bar{x}(0) - \bar{x}(p_{k+1}) \| \leq (d+e) \| p_{k+1} \| \quad , \quad \text{and}$$

$$\| s_k \| := \| x_{k+1}-x_k \| \leq \| x_{k+1}-x^* \| + \| x_k-x^* \| \text{ yield}$$

$$\frac{\| p_{k+1} \|}{\| s_k \|} \geq (d+e)^{-1} \frac{\| x^*-x_{k+1} \|}{\| s_k \|} \geq (d+e)^{-1} \frac{\| x^*-x_{k+1} \|}{\| x^*-x_{k+1} \| + \| x^*-x_k \|}$$

Defining $r_k := \dfrac{\| x^*-x_{k+1} \|}{\| x^*-x_k \|}$, we have

$$\frac{\| p_{k+1} \|}{\| s_k \|} \geq (d+e)^{-1} \frac{r_k}{1+r_k}$$

Hence $$0 = \lim_{k \to \infty} \frac{\| p_{k+1} \|}{\| s_k \|} \geq (d+e)^{-1} \lim_{k \to \infty} \frac{r_k}{1+r_k},$$ which implies

$\lim_{k \to \infty} r_k$. Hence, $\{x_k\}$ converges Q-superlinearly, as was to be shown. This completes the proof of the theorem.

5. Computational Results.

The following 4×4 nonlinear complementarity problem is given in Kojima [11].

$$f(x) := \begin{bmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{bmatrix} \geqq 0 ,$$

$$x := (x_1, x_2, x_3, x_4) \geqq 0 , \quad \langle x, f(x) \rangle = 0 .$$

The unique solution is given as

$$\hat{x}_1 = \sqrt{6}/2 = 1.2247449 , \ \hat{x}_2 = 0, \ \hat{x}_3 = 0 , \ \hat{x}_4 = 0.5 .$$

The generalized equation representing this nonlinear complementarity problem is $0 \in f(x) + N_{R_+^4}(x)$. This generalized equation has the linearization at $\bar{x}$ given by

$$0 \in f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + N_{R_+^4}(x) .$$

A quasi-Newton method replaces $f'(\bar{x})$ with an approximation B, resulting in the generalized equation

$$0 \in f(\bar{x}) + B\cdot(x-\bar{x}) + N_{R_+^4}(x) ,$$

which corresponds to the linear complementarity problem

$$x \geqq 0 , \quad f(\bar{x}) - B\bar{x} + Bx \geqq 0, \quad \langle x, f(\bar{x}) - B\bar{x} + Bx \rangle = 0 .$$

The update function used for this test case is Broyden's rank-one update, that is, given iterates x_n and x_{n+1}, and current approximation B_n, the new approximation B_{n+1} to $f'(x_{n+1})$ is given by

$$B_{n+1} := B_n + \frac{(y_n - B_n S_n)S_n^T}{\langle S_n, S_n \rangle} \text{ , where } S_n := x_{n+1} - x_n$$

and $y_n := f(x_{n+1}) - f(x_n)$.

We note that if $S_n = 0$, then $x_{n+1} = x_n$. Since, by definition, $0 \in f(x_n) + B_n(x_{n+1} - x_n) + N_{R_+^4}(x_{n+1})$, replacing x_{n+1} by x_n yields $0 \in f(x_n) + N_{R_+^4}(x_n)$, so that x_n is a solution to the original nonlinear complementarity problem if $S_n = 0$, and the iterative procedure is terminated.

The results of applying Lemke's algorithm to the linear complementarity problems with Broyden rank one approximations to the derivative f' are given in Table 1.

Table 1 Kojima Example

Initial x	Iterations	Final x	Final f
1.	6	(1.2247449,0,0,.49999993)	($.2.10^{-6}$, 3.224745, 5.0000002, $-.9.10^{-7}$)
3.	11	(1.2247449,0,0,.49999995)	(0, 3.224745, 5.0000000, $-.9.10^{-7}$)
5.	12	(1.2247449,0,0,.49999999)	($-.2.10^{-6}$,3.224745, 4.9999998, $-.9.10^{-7}$)

Three starting points, with each component of x initialized to the entry in column 1, all lead to a convergent sequence of iterates. The number of iterations is listed in column 2, with the final value of x and f(x) given in columns 3 and 4, respectively. Each iteration consists of Lemke's algorithm applied to a 4×4 linear complementarity problem, each of which required two pivot operations of Lemke's algorithm to solve.

ACKNOWLEDGEMENTS

Parts of the reasearch reported in this paper are based on the author's doctoral thesis submitted to the Department of Industrial Engineering at the University of Wisconsin-Madison, written under the direction of Professor Stephen M. Robinson.

I wish to express my sincere gratitude to Professor Stephen M. Robinson for his advice and support. His continued encouragement, guidance and understanding throughout my graduate studies have been invaluable to me.

REFERENCES

[1] Broyden, C. G., Dennis, J. E. and Moré, J. (1973). On the Local and Superlinear Convergence of Quasi-Newton Methods. J. Inst. Math. Appl. 12, 223-246.

[2] Brodlie, K. W. (1977). Unconstrained Minimization. In: The State of the Art in Numerical Analysis, ed. D. Jacobs, pp. 229-268. Academic Press, New York.

[3] Cottle, R. W. and Dantzig, G. B. (1967). Complementary Pivot Theory of Mathematical Programming. Technical Report 67-2, Operations Research House, Stanford University, California.

[4] Dennis, J. E. and Moré, J. (1974). A Characterization of Superlinear Convergence and its Application to Quasi-Newton Methods. Math. Comp. 28, 549-560.

[5] Dennis, J. E. and Moré, J. (1977). Quasi-Newton Methods: Motivation and Theory. SIAM Review, 19 , 46-89.

[6] Dennis, J. E. and Schnabel, J. (1978). Least Change Secant Updates for Quasi-Newton Methods. Technical Report TR 78-344, Computer Science Department, Cornell University, New York.

[7] Gill, P. E., Golub, G., Murray, W. and Saunders, M. (1974). Methods for Modifying Matrix Factorizations. Math. Comp. 28, 505-536.

[8] Gill, P. E. and Murray, W. (eds) (1974). Numerical Methods for Constrained Optimization. Academic Press, New York.

[9] Han, S. -P. (1976). Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems. Math. Prog. 11, 263-282.

[10] Josephy, N. H. (1979). Newton's Method for Generalized Equations. Technical Summary Report #1965, Mathematics Research Center, University of Wisconsin-Madison.

[11] Kojima, M. (1974). Computational Methods for Solving the Nonlinear Complementarity Problem. Keio Engineering Reports, 27, 1-41.

[12] Ortega, J. M. and Rheinboldt, W. C. (1970). Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York.

[13] Robinson, S. M. (1976a). An Implicit-Function Theorem for Generalized Variational Inequalities. Technical Summary Report No. 1672, Technical Summary Report #1672, Mathematics Research Center, University of Wisconsin-Madison.

[14] Robinson, S. M. (1976b). Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems. SIAM J. Num. Anal. 13, 497-513.

[15] Robinson, S. M. (1976c). Regularity and Stability for Convex Multivalued Functions. Math. of Operations Research 1 , 130-143.

[16] Robinson, S. M. (1978). Strongly Regular Generalized Equations. Technical Summary Report #1877, Mathematics Research Center, University of Wisconsin-Madison.

[17] Robinson, S. M. (1979). Generalized Equations and Their Solutions, Part I: Basic Theory. Math. Programming Studies 10 , 128-141.

[18] Tapia, R. A. (1977). Diagonalized Multiplier Methods and Quasi-Newton Methods for Constrained Optimization. J. Optimization Theory Appl. 22 , 135-194.

NHJ:db

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 1966	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) QUASI-NEWTON METHODS FOR GENERALIZED EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Norman H. Josephy		8. CONTRACT OR GRANT NUMBER(s) DCR 74-20584 DAAG29-75-C-0024 MCS 74-20584 A02
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 5 - Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE June 1979 13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) MRC-TSR-1966		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

Technical Summary Rept.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Variational Inequalities, Complementarity, Quasi-Newton Methods.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Newton's method is a well known and often applied technique for computing a zero of a nonlinear function. Situations arise in which it is undesirable to evaluate, at each iteration, the derivative appearing in the Newton iteration formula. In these cases, a technique of much modern interest is the quasi-Newton method, in which an approximation to the derivative is used in place of the derivative. By using the theory of generalized equations, quasi-Newton methods are developed to solve problems arising in both mathematical programming and mathematical economics.

(continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

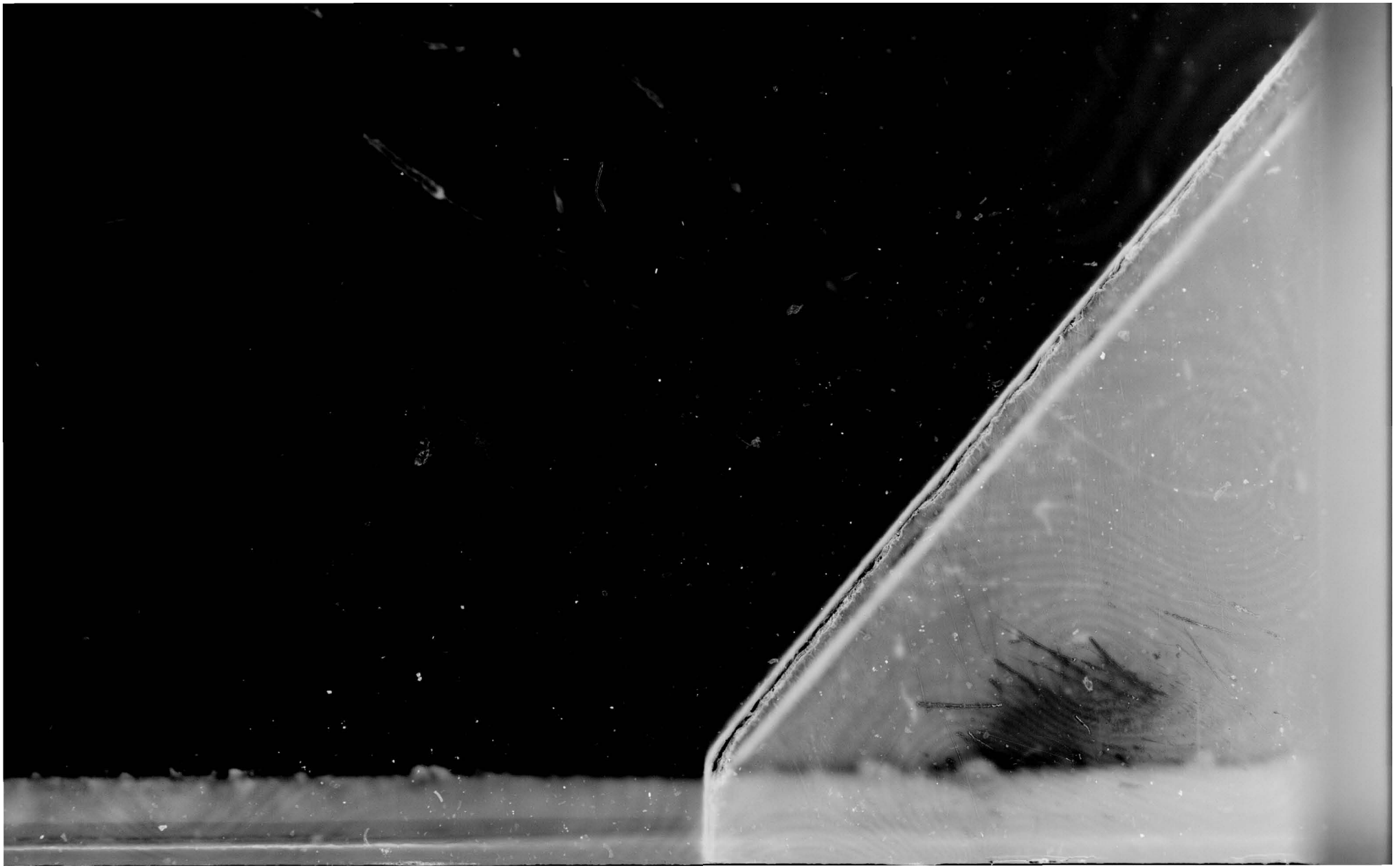
UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

221 200

JOB

Abstract (continued)

We prove two results concerning the convergence and convergence rate of quasi-Newton methods for generalized equations.

We present computational results of quasi-Newton methods applied to a nonlinear complementarity problem of Kojima [11].